

Digital Watermarks For Checking Authenticity of Printed Objects

Related Application Data

The subject matter of the present application is related to that disclosed in US Patent Applications 09/840,016, filed April 20, 2001, 60/263,987, filed January 24, 2001, 5 09/689,226, filed October 11, 2000, and 09/234,780, filed January 20, 1999, which are hereby incorporated by reference.

Technical Field

The invention relates to steganography and data hiding, and specifically relates to the use of digital watermarks for authentication of printed objects.

Background and Summary

Digital watermarking is a process for modifying physical or electronic media to embed a machine-readable code into the media. The media may be modified such that the embedded code is imperceptible or nearly imperceptible to the user, yet may be detected through an automated detection process. Most commonly, digital watermarking 15 is applied to media signals such as images, audio signals, and video signals. However, it may also be applied to other types of media objects, including text documents (e.g., through line, word or character shifting), software, multi-dimensional graphics models, and surface textures of objects.

Digital watermarking systems typically have two primary components: an 20 encoder that embeds the watermark in a host media signal, and a decoder that detects and automatically reads the embedded watermark from a signal suspected of containing a watermark (a suspect signal). The encoder embeds a watermark by altering the host media signal. The reading component analyzes a suspect signal to detect whether a watermark is present. In applications where the watermark encodes information, the 25 reader extracts this information from the detected watermark.

Several particular watermarking techniques have been developed. The reader is presumed to be familiar with the literature in this field. Particular techniques for embedding and detecting imperceptible watermarks in media signals are detailed in the assignee's co-pending application serial number 09/503,881 and US Patent 6,122,403, which are hereby incorporated by reference.

The invention provides methods for embedding digital watermarks for authentication of printed objects, and corresponding methods for authenticating these objects. One aspect of the invention is a method of embedding a digital watermark in a digital image to be printed on an object. The method embeds an auxiliary signal in the digital image so that the auxiliary signal is substantially imperceptible, yet machine readable. It converts the image to a halftone image using an inherently unstable halftone screen structure that is likely to cause ink flow errors when reproduced. The errors introduced by reproducing the unstable halftone screen structure incorrectly are automatically detectable by reading the auxiliary signal.

Another aspect of the invention is a method of authenticating a printed object. The method receives a digital image scanned of the printed object, and detects a digital watermark in the digital image. It then uses a measurement of strength of the digital watermark to detect reproduction errors due to inaccurate reproduction of unstable halftone screen structures in the printed object. Another aspect of the invention is a method of authenticating a printed object using different watermarks in the object. This method detects a first digital watermark from a substrate of the printed object, and detects a second digital watermark from an image scanned from the printed object. It uses the relationship between the first and second digital watermarks to determine authenticity of the printed object.

Another aspect of the invention is a method of authenticating a printed object using the geometric relationship between visible fiducials and embedded watermarks on a printed object. This method detects a visible fiducial in an image scanned from the printed object, and detects a location of a digital watermark hidden in the printed object. It determines authenticity of the printed object by comparing the location of the digital

watermark to the visible fiducial. Further features will become apparent with reference to the following detailed description and accompanying drawings.

Brief Description of the Drawings

5 Fig. 1 illustrates a flow diagram of watermark embedding and halftone screening process using inherently unstable screens to create an authentication watermark on a printed object.

 Fig. 2 illustrates an example showing different halftone structures, including a relatively stable structure (A), and inherently unstable structures (B and C).

10 Fig. 3 is a diagram illustrating a process for detecting an authentication watermark from a scanned image of the watermarked object, and for assessing the validity of the watermark signal.

Detailed Description

Using Inherently Unstable Structure as Screening Elements

15 Fig. 1 illustrates a flow diagram of watermark embedding and halftone screening process using inherently unstable screens to create an authentication watermark on a printed object. This process generates a digital watermark signal and embeds it in a halftone image. The screens used to create the halftone image have inherently unstable dot structures that are difficult to reproduce. As such, if the printed object is copied (e.g.,
20 scanned and re-printed) the dot structures are likely to change. A watermark reader processes a digital image of the object, and detects errors in the watermark signal due to the changes in the ink structures. The extent of these errors indicate the extent to which the printed object is likely to be a counterfeit.

 Turning to Fig. 1, this example digital watermark embedding process starts with a
25 message (100) to be hidden in the printed document. The message comprises a set of binary (or M-ary) symbols, including both fixed symbols to assist in detection and message interpretation (e.g., a message common to a particular message protocol), and

variable symbols to carry a variable message (e.g., a unique message for a particular printed item). The message is error correction encoded (e.g., convolution, turbo, BCH, repetition coding)(102), and then spread spectrum modulated with a carrier signal (104). In this example, the carrier signal comprises a pseudorandom image signal that

5 corresponds to a block of image samples in the host image. The spread spectrum modulation function may be implemented using a variety of spreading functions, such as convolving, multiplication, XOR, XNOR operations between the message signal and the carrier to create a watermark signal. Preferably, the watermark signal has a calibration signal component, e.g., signal patterns such as peaks in the autocorrelation or Fourier
10 domain that facilitate detection, and calculation of rotation, scale and translation parameters of the watermark signal in an image scan of a printed object carrying the watermark. Examples of spread spectrum modulation and calibration signals are disclosed in co-pending application serial number 09/503,881 and US Patent 6,122,403, incorporated above.

15 In this implementation, a digital watermark signal is generated at a source resolution that is lower than the target resolution of the half tone dots used to create the printed image. The watermark signal is embedded using an embedding process (106) that adds samples of the watermark signal to corresponding samples of the host image (108). The embedding function may also include perceptual masking of the watermark signal,
20 such as adjusting the watermark signal strength according to Human Visibility System modeling filters that determine the capacity of the host image to hide the watermark signal. The embedding process (106) creates a watermarked image signal comprising an array of multi-level per pixel values at specified resolution.

Next, the watermarked image signal is converted to an image comprised of
25 halftone dots at a target resolution that is higher than the source resolution of the watermarked signal (110). The half tone dot represents the presence or absence of ink at a particular spatial location on a printed object. The halftone process uses inherently unstable screens (112) to convert selected multi-level per pixel samples in the

watermarked image to corresponding halftone dot structures comprising an array of halftone dots.

In some applications, like corporate logos, graphic symbols, etc. the host image can be particularly adapted to carry a watermark signal. In other words, the watermark signal and corresponding image are designed together to form the watermarked image. In this case, there is no need for a generic watermark embedding technique that is broadly applicable to a variety of images. Also, this enables the image to be designed with special features that assist in watermark detection, like visible fiducials that enable the watermark reader to align the image geometrically before detecting the watermark or performing other image analysis. Conventional pattern recognition and registration routines may be used to align the image as a pre-process step before authentication.

Fig. 2 illustrates an example showing different halftone structures, including a relatively stable structure (A), and inherently unstable structures (B and C). To create an authentic printed object, these unstable structures are used to create a halftone image, which is then printed with a high quality printer (110) capable of accurately retaining the unstable structure of ink on the printed object. Later, if the printed object is scanned and printed, it is difficult to reproduce the structures on typical ink jet printers, even if the authentic object is scanned with a high resolution scanner. For such ink jet printers, an inherently unstable structure creates the likelihood of ink flow such that ink jet printer is unlikely to reproduce the structure with the same ink coverage (e.g. the ink density over the same area changes causing a change in color or luminance of the pixel). The ink flow at the boundary of the unstable structure causes more area to be covered, and as such, an ink jet printer with such ink flow cannot accurately reproduce the unstable structure without changing the density of ink over the area covered by the structure in the authentic printed object.

As an example, consider the case where the watermarked signal has a resolution of 100 dpi and the halftone dot resolution is 500 dpi. In other words, each 100 dpi pixel in the watermarked signal corresponds to a 5 by 5 array of halftone dots. To represent a multilevel per pixel value of 8/25, any of the structures A-C could be used, but structures

B and C are more likely to cause ink flow problems if a counterfeiter tries to reproduce them using an ink jet printer with ink flow problems. For example, the same 5 by 5 area may be reproduced as 10/25 (10 halftone dots with ink) instead of 8/25 (8 halftone dots with ink).

5 Traditional halftone screening algorithms avoid square clusters of halftone dots in cases where about 50% of the dots are turned on. The reason for avoiding this pattern of halftone dot clusters is that it causes ink to flow between the dots where the corners of the square clusters meet. Structure D in Fig. 2 illustrates an example of an unstable screen with square clusters joined at the corners. While avoided in the traditional case, this
10 arrangement may be used advantageously to authenticate a printed object because certain printing presses can represent the pattern accurately on paper. However, when reproduced by the counterfeiter on a different printer, error due to ink flow are likely.

 The error in reproduction can cause the watermark to be un-detectable by an automated reader, or may create enough errors in the watermark signal to indicate that the
15 printed object is likely a fake. In the first case, if the watermark reader cannot detect the watermark, the printed object is deemed not to be authentic. In the second case, the authenticity is determined by evaluating the watermark strength relative to a threshold. This process embeds a watermark at a sufficiently low resolution and sufficiently high strength to enable the watermark to be read using commercial scanners and web cameras,
20 and it also enables detection of invalid printed objects.

 Fig. 3 is a diagram illustrating a process for detecting an authentication watermark from a scanned image of the digitally watermarked object, and for assessing the validity of the watermark signal. First, the user holds the printed object up to a web camera or places it on an image scanner plate. The watermark reader receives a digital image (120)
25 of the printed object, and pre-processes it to prepare for watermark detection (122).

 The pre-processing (122) entails a conversion to the color plane in which the watermark signal is embedded (e.g., blue channel, yellow channel, luminance, chrominance, etc.), and a filtering of blocks of the image to identify candidate blocks that are likely to have the authentication watermark. Examples of the filtering include

looking for portions of the image with a certain signal activity (texture), color, edge concentration, etc.

Next, the watermark detector proceeds to detect the watermark signal and determine orientation parameters of watermark signal blocks (e.g., rotation, scale, translation). The candidate image blocks identified in the pre-processing stage are transformed into a transform domain for detection of the calibration component. For example, candidate blocks are transformed into an autocorrelation domain or Fourier domain, where peaks of the calibration component are detected and correlated with a reference signal having attributes of the known calibration signal component. An example of this technique using the Fourier-Mellin transform to recover rotation and scale is described in co-pending application serial number 09/503,881 and US Patent 6,122,403, incorporated above. This example detector correlates the pseudorandom phase information of the calibration signal with phase information of the received image, after compensating for rotation and scale to get the coordinates of the origin of a watermark block.

Next, the watermark detector reads the message embedded in the watermark signal (128). In this implementation, the detector is calibrated to read the watermark at its original resolution at the time of embedding (e.g., 100 dpi in the example above). To reduce the interference of the original host signal, the detector predicts the original signal and subtracts it, leaving an estimate of the watermark signal. The detector then performs the inverse of the spread spectrum modulation to compute raw estimates of the error correction encoded message symbols. These are soft estimates weighted by a probability reflecting the statistical confidence in the accuracy of the symbol. The soft estimates are supplied to an error correction decoder compatible with the one in the encoder, which generates a message, including the fixed and variable message symbols.

Error detection symbols, such as CRC bits computed on the message, indicate whether the decoded message is valid. If the detector is unable to extract a valid watermark, the printed object is deemed to be a copy. A successful watermark detection is measured by the degree of correlation between the detected signal and the reference

signal in the calibration phase, and/or a valid message as determined by the error detection evaluation.

If a valid watermark message is recovered, the detector proceeds to measure the strength of the watermark signal. There are multiple metrics for assessing watermark strength, including the degree of correlation between the reference signal and the detected signal, and a measure of symbol errors in the raw message estimates. One way to measure the symbol errors is to reconstruct the raw message sequence using the same error correction coding process of the embedder on the valid message extracted from the watermark. This process yields, for example, a string of 1000 binary symbols, which can be compared with the binary symbols estimated at the output of the spread spectrum demodulator. The stronger the agreement between the reconstructed and detected message, the stronger the watermark signal.

More specifically, an approach for measuring the strength of the watermark signal is as follows:

1. Use the message payload read from the watermark to re-create the original embedded bit sequence (including redundantly encoded bits from error correction coding) used for the watermark.
2. Convert the original bit sequence so that a zero is represented by -1 and a one is represented by 1.
3. Multiply (element-wise) the soft-valued bit sequence used to decode the watermark by the sequence of step 2.
4. Create one or more measures of watermark strength from the sequence resulting in the previous step. One such measure is the sum of the squares of the values in the sequence. Another measure is the square of the sum of the values in the sequence. Other measurements are possible as well. For example, soft bits associated with high frequency components of the watermark signal may be analyzed to get a strength measure attributed to high frequency components. Such high frequencies are likely to be more sensitive to

degradation due to photocopying, digital to analog and analog to digital conversion, scanning and re-printing, etc.

5. Compare the strength measures to thresholds to decide if the suspect image has been captured from an original or a copy of the printed object. The threshold is derived by evaluating the difference in measured watermark strength of copied vs. original printed objects on the subject printer platform used to create the original, and a variety of copiers, scanners and printers used to create copies.

This same technique of measuring symbol errors can be applied to two or more different watermarks embedded at different spatial resolutions. Each of the watermarks may have the same or different message payloads. In the first case where the watermarks have the same message payloads, the message extracted from one of the watermarks may be used to measure bit errors in each of the other watermarks. For example, the message payload from a robust watermark embedded at a low spatial resolution may be used to measure the bit errors from a less robust watermark at a higher spatial resolution. If the watermarks carry different message payloads, then error detection bits, such as CRC bits, can be used in each message payload to ensure that the message is accurately decoded before re-creating the original, embedded bit sequence.

Using two or more different watermarks enables a threshold to be set based on the ratio of the signal strength of the watermarks relative to each other. In particular, the signal strength of a first watermark at a high resolution (600-1200 dpi) is divided by the signal strength of a second watermark at a lower resolution (75-100 dpi). In each case, the signal strength is measured using a measure of symbol errors or some other measure (e.g., correlation measure).

If the measured strength exceeds a threshold, the detector deems the watermark signal to be authentic and generates an authentication signal (132). This signal may be a simple binary value indicating whether or not the object is authentic, or a more complex image signal indicating where bit errors were detected in the scanned image.

The watermark and host signal can be particularly tailored to detect copying by photo duplication and printing/re-scanning of the printed object. This entails embedding the watermark with selected screening structures at particular spatial frequencies/resolutions that are likely to generate message symbol errors when the object is re-printed. This detection process has an additional advantages in that it enables automatic authentication, it can be used with lower quality camera devices such as web cams and common image scanners, and it allows the watermark to serve the functions of determining authenticity as well as carrying a message payload useful for a variety of applications.

The message payload can include an identifier or index to a database that stores information about the object or a link to a network resource (e.g., a web page on the Internet). The payload may also include a covert trace identifier associated with a particular authentic item, batch of items, printer, or distributor. This enables a counterfeit object, or authentic object that has been printed without authority to be detected and traced to a particular source, such as its printer, distributor or batch number.

The payload may also carry printer characteristics or printer type information that enables the watermark reader to adapt its detection routines to printer types that generated the authentic object. For example, the payload may carry an identifier that specifies the type of halftoning used to create the authentic image, and more specifically, the attributes of the halftone screen. With this information, the reader can check authenticity by determining whether features associated with the halftone screen exist in the printed object. Similarly, the reader can check for halftone screen attributes that indicate that a different halftone screen process has been used (e.g., a counterfeit has been created using a different halftone screen). One specific example is a payload that identifies the halftone screen type and paper type. The reader extracts this payload from a robust watermark payload and then analyzes the halftone screen and paper attributes to see if they match the halftone type and paper type indicated in the watermark payload. For example, the halftone type can specify the type of unstable screen used to create an authentic image. If

this unstable screen is not detected (e.g., by absence of a watermark embedded in the unstable screen), then the image is considered to be a fake.

A related approach for analyzing halftone type is to look for halftone attributes, like tell-tale signs of stochastic halftone screens vs. ordered dither matrix type screens.

- 5 Dither matrix screens used in low end printers tend to generate tell tale patterns, such as a pattern of peaks in the Fourier domain that differentiate the halftone process from a stochastic screen, such as an error diffusion process, which does not generate such tell-tale peaks. If the reader finds peaks where none were anticipated, then the image is deemed a fake. Likewise, if the reader finds no peaks where peaks were anticipated, then
- 10 the image is also deemed a fake. Before performing such analysis, it is preferable to use the embedded digital watermark to re-align the image to its original orientation at the time of printing. Attributes due to the halftone screen can then be evaluated in a proper spatial frame of reference. For example, if the original ordered dither matrix printer created an array of peaks in the Fourier domain, then the peak locations can be checked
- 15 more accurately after the image is realigned.

Additional Authentication Methods

- Another way to authenticate the printed object is to embed separate digital watermarks in the object, including one in the substrate of the object, and another hidden in a visible image printed on the substrate. A verification device looks for these
- 20 watermarks in a scanned image of the object, and used the expected relationship between them to check authenticity. In one case, the presence of one watermark without the other indicates that the object is a fake. For example, if the image on an authentic object is scanned and re-printed on another substrate of a fake object, then the new substrate will not have the watermark embedded in the substrate.

- 25 In one implementation, the substrate carries a digital watermark in the form of the calibration signal as described above. This digital watermark is embedded so as to be substantially imperceptible by modulating the surface micro-topology of the substrate to form the calibration signal, which is a pseudorandom pattern with registration peaks in a transform domain (e.g., autocorrelation or Fourier transform domain). In the

detection/verification process, a watermark detector processes a scanned image of the object, looking for the calibration signal. If none is present, or only a weak signal as measured by the extent of correlation with the known calibration signal, then the object is deemed a fake.

5 One form of substrate watermark is a layer of UV ink or lacquer that modulates the topology of the object surface to form the substrate watermark signal. This UV ink is transparent or semi-transparent, yet forms a surface topology that modulates the luminance of the object surface to embed the watermark signal. This layer of ink can be formed on the surface of the object by printing an image of the substrate watermark onto
10 the object using the UV ink or lacquer.

The other watermark may be embedded in a luminance or other color channel of a color image or gray-scale image printed on the object. For the sake of explanation, let's refer to this other watermark as the image watermark, as opposed to the substrate watermark.

15 For additional verification, the substrate and image watermarks can carry payload information that must satisfy a predetermined relationship to establish the authenticity of the object bearing these digital watermarks. For example, the message payloads may need to satisfy a predetermined relationship (e.g., one is a cryptographic transform of the other).

20 Either the substrate or the image watermark may also be embedded at predetermined locations relative to the other watermark or some visible fiducial on the object that can be detected automatically. Such a geometric relationship between the digital watermark and a fiducial, or between different digital watermarks is designed to be difficult to re-produce accurately in a copy of the object. To check authenticity, the
25 watermark location is extracted and checked relative to the location of a second watermark or fiducial. If the relative positions of the watermark and fiducial/second watermark do not fall within a predetermined tolerance, the object is deemed to be a fake. The fiducial may be a simple visible pattern detectable by pattern recognition techniques,

the edge of page, the border of text (e.g., margin of the page), or some other distinguishable structure.

Concluding Remarks

5 Having described and illustrated the principles of the technology with reference to specific implementations, it will be recognized that the technology can be implemented in many other, different, forms. To provide a comprehensive disclosure without unduly lengthening the specification, applicants incorporate by reference the patents and patent applications referenced above.

10 The methods, processes, and systems described above may be implemented in hardware, software or a combination of hardware and software. For example, the auxiliary data encoding processes may be implemented in a programmable computer or a special purpose digital circuit. Similarly, auxiliary data decoding may be implemented in software, firmware, hardware, or combinations of software, firmware and hardware. The
15 methods and processes described above may be implemented in programs executed from a system's memory (a computer readable medium, such as an electronic, optical or magnetic storage device).

20 The particular combinations of elements and features in the above-detailed embodiments are exemplary only; the interchanging and substitution of these teachings with other teachings in this and the incorporated-by-reference patents/applications are also contemplated.